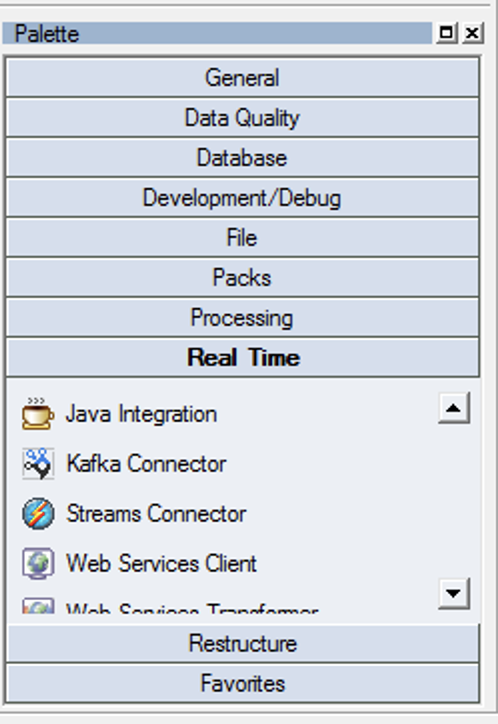
# Вывод табличных данных в XML средствами IBM DataStage с дополнительной обработкой в Java-операторе

<https://github.com/zinal/dssamples/tree/main/xml-output/complex>

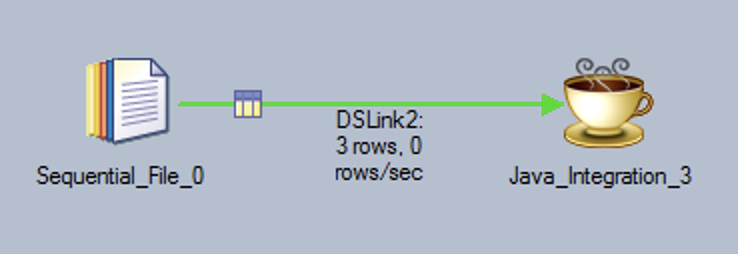
В разделе «Real Time» есть оператор «Java Integration» (см. скриншот ниже).



Он позволяет выполнять произвольные виды обработки записей из одного либо нескольких входных потоков, а также возвращать записи в один или несколько выходных потоков.

В нашем простейшем примере у Java-оператора не будет выходных потоков, результаты своей работы он будет записывать в файл.

Простейшее задание читает CSV и подаёт на обработку в Java-оператор:



При чтении CSV необходимо определить структуру полей, как показано в примере настроек ниже:

Изображение выглядит как стол

Автоматически созданное описание

Если предполагаются пустые значения (и их надо обрабатывать как NULL), важно выставить свойство «Null field value» в значение '' (две одинарные кавычки – пустое текстовое значение):

Изображение выглядит как текст

Автоматически созданное описание

Логика вывода данных определяется кодом оператора на языке Java. Например:

Изображение выглядит как текст

Автоматически созданное описание

В настройках оператора необходимо определить путь к скомпилированному JAR-файлу, имя класса реализации оператора, объём используемой памяти.

Рекомендуется также установить параметр «-Xdump:none» для отключения автоматического формирования диагностических дампов при ошибках выполнения Java-оператора:

Изображение выглядит как стол

Автоматически созданное описание

У Java-оператора могут быть свои собственные настройки, указываемые в разделе «Custom properties». В данном случае имя выходного файла определяется через значение параметра задания DataStage.